Confidence Intervals for Treatment Effect in Meta-Analysis

Andrew L. Rukhin[1]
Statistical Engineering Division
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

Abstract

Confidence intervals for the treatment effect in random effects meta-analysis model obtained from Harville-Jeske-Kenward-Roger approach are obtained in explicit form. They are compared to some other intervals commonly used in collaborative studies with a small number of participants possibly with heterogeneous, study-specific variances. Monte Carlo simulation experiments recommend the latter intervals.

Keywords: Fisher information; Heteroscedasticity; Interlaboratory study; Random effects model; Restricted likelihood.

1   Confidence estimation in meta-analysis problems

The subject of interest here is confidence intervals for the common mean when several different studies, methods, instruments or laboratories measure a given property of the same material or the difference between two treatements. Combination or pooling of such measurements to allow statistical analysis of several individual studies is a goal of meta-analysis. Although some debate concerning advantages of random effects models in meta-analysis continues, (see Borenstein et al., 2009), the following heterogeneous model has become a common tool of choice.

Denote by n_i the number of observations made in the laboratory i, $i = 1, \ldots, p$. In the interlaboratory studies applications which are of interest here, neither p nor n_i are large, but a Gaussian distribution condition is

---

[1]E-mail address: andrew.rukhin@nist.gov

made. Namely, the observed data $x_{ik}, k = 1, \ldots, n_i$, is assumed to have the form

$$x_{ik} = \mu + \ell_i + \epsilon_{ik}, \tag{1}$$

where μ is the treatment effect or the property value, ℓ_i represents the study (or method) effect, which is normal with mean 0 and unknown variance σ^2. The independent normal zero mean random errors ϵ_{ik}, have unknown (different) variances τ_i^2. For a fixed i, $x_i = \sum_k x_{ik}/n_i$, is normally distributed with the mean μ and the variance $\sigma^2 + \sigma_i^2$, where $\sigma_i^2 = \tau_i^2/n_i$. If $\sigma^2 + \sigma_i^2$ were known up to a factor, then the least squares estimator of μ could be used, $\tilde{\mu} = \sum_i \omega_i x_i$ with normalized weights

$$\omega_i = \frac{\frac{1}{\sigma^2+\sigma_i^2}}{\sum_j \frac{1}{\sigma^2+\sigma_j^2}}. \tag{2}$$

In this situation,

$$\mathrm{Var}(\tilde{\mu}) = \Phi = \left[\sum_i \frac{1}{\sigma^2 + \sigma_i^2}\right]^{-1}. \tag{3}$$

Since the variances are unknown, these optimal weights have to be estimated. In some problems of meta-analysis the sample sizes n_i are not available, but when they are, the classical unbiased statistic $s_i^2 = \sum_j (x_{ij} - x_i)^2/[n_i(n_i-1)]$ has the distribution $\sigma_i^2 \chi^2(\nu_i)/\nu_i, \nu_i = n_i - 1$, and is independent of x_i and $s_j^2, j \neq i$. This is the situation studied in this paper.

To estimate σ's the restricted maximum likelihood estimator (REML) is commonly employed. It is well known that the plug-in version of (3), which replaces the unknown $\sigma^2, \sigma_1^2, \ldots, \sigma_p^2$ by statistics $\tilde{\sigma}^2, \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2$ such that $E(\tilde{\sigma}^2 + \tilde{\sigma}_i^2) \leq \sigma^2 + \sigma_i^2$, underestimates the variance of the corresponding common mean estimator. Since (3) is an increasing function of σ's, positively biased estimators partly compensate for this inequality. Our goal is to derive REML based confidence intervals for the treatment effect in model (1) which includes corrections to the traditional method by using Harville-Jeske-Kenward-Roger approach.

The organization of this paper is as follows. In the next section 2 the method of Harville and Jeske (1992), Kenward and Roger (1997) to obtain confidence intervals is discussed. Explicit formulas for all characteristics which determine these intervals are found. The common methods of standard error evaluation for the treatment effect μ do not explicitly take into account the sample sizes n_i. When the REML variance estimator ia applied to (7), the resulting standard error of the confidence intervals exhibited in

2

section 3, depends on the degrees of freedom ν_i. These confidence intervals are compared via a Monte Carlo study in section 4. All mathematical derivations are collected in the Appendix.

2 Restricted maximum likelihood method: variance approximations and information matrix

In a general context of mixed effects linear models, Harville and Jeske (1992), Sec 4.2, suggested an estimator of the variance of a sample counterpart of the least squares statistic, which in our case is the weighted average,

$$\tilde{x} = \sum_i x_i \tilde{\omega}_i = \tilde{\Phi} \sum_i \frac{x_i}{\tilde{\sigma}^2 + \tilde{\sigma}_i^2}$$

with $\tilde{\omega}_i = (\tilde{\sigma}^2 + \tilde{\sigma}_i^2)^{-1}/(\sum_j (\tilde{\sigma}^2 + \tilde{\sigma}_j^2)^{-1})$. They suggested to use the REML variances estimator $\tilde{\sigma}^2, \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2$ and put forward two following approximations based on Taylor's formula or the propagation-of-error method. The first one deals with the mean squared difference between $\tilde{x}$ and $\tilde{\mu}$:

$$E(\tilde{x} - \tilde{\mu})^2 \approx \text{tr}\,(\mathcal{V}\Lambda). \tag{4}$$

Here $\mathcal{V}$ is the mean squared error matrix of $(\tilde{\sigma}^2, \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2)$, and Λ is the covariance matrix of the vector $(\tilde{\mu}_0', \tilde{\mu}_1', \ldots, \tilde{\mu}_p')^T, \partial \sigma_0^2 = \partial \sigma^2$. This approximation was originally introduced by Kackar and Harville (1984).

The second approximation corrects for the bias of the plug-in estimator $\tilde{\Phi}$ of Φ,

$$E\tilde{\Phi} \approx \Phi + \frac{1}{2}\text{tr}\,(\mathcal{V}H) = \Phi - \text{tr}\,(\mathcal{V}\Lambda), \tag{5}$$

where H is the Hessian of Φ, which is a negative semidefinite matrix, evaluated at $\sigma^2, \sigma_1^2, \ldots, \sigma_p^2$. In the model (1), $\Lambda = -H/2$. The formula (5) requires (approximate) unbiasedness of $(\tilde{\sigma}^2, \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2)$, so that the linear term in $(\tilde{\sigma}^2 - \sigma^2, \tilde{\sigma}_1^2 - \sigma_1^2, \ldots, \tilde{\sigma}_p^2 - \sigma_p^2)$ can be neglected. Since $\tilde{x}$ is known to be an unbiased estimator of μ, such that $\tilde{x} - \tilde{\mu}$ is independent of $\tilde{\mu}$, one gets $Var(\tilde{x}) = E(\tilde{x} - \tilde{\mu})^2 + \Phi$, which suggests the formula,

$$\widetilde{Var}(\tilde{x}) = \tilde{\Phi} + 2\text{tr}(\tilde{\mathcal{V}}\tilde{\Lambda}) \tag{6}$$

Here $\tilde{\mathcal{V}}$ is the estimated mean squared error matrix of $(\tilde{\sigma}^2, \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2)$ and $\tilde{\Lambda}$ has a similar meaning.

Kenward and Roger (1997) gave a formalization of these approximations in a general mixed effects linear model when the inverse of the restricted

3

likelihood information matrix $\mathcal{J}$ is used in lieu of $\mathcal{V}$. Via a Monte Carlo study they demonstrated good performance of the resulting variance estimators and test statistics in several more general than (1) random effects models. The SAS procedure "MIXED" employs estimators and confidence intervals derived by this method.

Since matrices $\mathcal{V}$ and $\Lambda \neq 0$ are positive semidefinite, $\text{tr}(\mathcal{V}\Lambda) > 0$, and (6) seems to confirm negative bias of the estimator $\tilde{\Phi}$. Although in our problem $(\tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_p^2)$ can be assumed to be an (approximately) unbiased estimator of $(\sigma_1^2, \ldots, \sigma_p^2)$, it is easy to verify that all non-negative estimators of σ^2 are biased. To adjust for that fact, the formula

$$E\tilde{\Phi} \approx \Phi + \Upsilon \Phi^2 \sum_i \frac{1}{(\sigma^2 + \sigma_i^2)^2} - \text{tr}\left(\mathcal{V}\Lambda\right),$$

could be put in place of (5). Thus only positively biased estimators of σ^2 have a chance to give rise to (approximately) unbiased estimates of Φ. Commonly $\tilde{\sigma}^2$ is a positive part of an unbiased estimator, so that its bias is positive, $\Upsilon = E(\tilde{\sigma}^2 - \sigma^2) > 0$, and it can be substantial especially for small σ^2. Kenward and Roger (2009) suggest a different bias correction term in general linear models obtained in terms of the inverse of the (full) likelihood information matrix. However since in our situation the covariance matrix of $(x_1, \ldots, x_p)$ is a linear function of the parameters $\sigma^2 + \sigma_i^2, i = 1, \ldots, p$, according to these authors, no bias correction in (6) is required.

When the estimates $\tilde{\sigma}^2$'s are substituted for the unknown σ^2's, the estimator (6) takes the form

$$\widetilde{Var}(\tilde{x}) = \tilde{\Phi} - \tilde{\Upsilon}\tilde{\Phi}^2 \sum_i \frac{1}{(\tilde{\sigma}^2 + \tilde{\sigma}_i^2)^2} + 2\text{tr}(\tilde{\mathcal{V}}\tilde{\Lambda}). \qquad (7)$$

By taking advantage of the specific form of the linear model (1), we give now very explicit formulas for $\mathcal{W}, \Lambda$ and $\widetilde{Var}(\tilde{x})$ in (7) when REML variance estimators are used.

The restricted likelihood function, when written as a function of sufficient statistics $x_i, s_i^2, i = 1, \ldots, p$, has the form

$$RL = -\frac{1}{2}\Big[\sum_i \frac{(x_i - \tilde{\mu})^2}{\sigma^2 + \sigma_i^2}$$

$$+ \log\left(\sum_i (\sigma^2 + \sigma_i^2)^{-1}\right) + \sum_i \log(\sigma_i^2 + \sigma^2) + \sum_i \frac{\nu_i s_i^2}{\sigma_i^2} + \sum_i \nu_i \log \sigma_i^2\Big]. \quad (8)$$

4

The inverse of the Fisher information matrix $\mathcal{J} = \mathcal{J}(\sigma^2, \sigma_1^2, \ldots, \sigma_p^2)$ based on (8) gives the asymptotic covariance matrix of the restricted maximum likelihood estimators which is used as $\mathcal{V}$.

For ω_i defined by (2) put

$$S_m = \sum_i \omega_i^m, \quad m = 2, \ldots,$$

The results in the Appendix, section 6.1, show that

$$\mathcal{J} = \begin{pmatrix} a_{00} & a^T \\ a & D + bb^T \end{pmatrix} \tag{9}$$

with

$$a_{00} = \frac{S_2 - 2S_3 + S_2^2}{2\Phi^2}, \tag{10}$$

a denoting the p-dimensional vector with coordinates a_i,

$$a_i = \frac{\omega_i^2(1 + S_2 - 2\omega_i)}{2\Phi^2}, \tag{11}$$

the vector b having coordinates b_i,

$$b_i = \frac{\omega_i^2}{\sqrt{2}\Phi}, \quad i = 1, \ldots, p, \tag{12}$$

and the diagonal matrix D given by elements d_{ii},

$$d_{ii} = \frac{\nu_i}{2\sigma_i^4} + \frac{\omega_i^2(1 - 2\omega_i)}{2\Phi^2}. \tag{13}$$

The inverse matrix $\mathcal{J}^{-1} = \mathcal{V}$ has the form,

$$\mathcal{V} = \begin{pmatrix} v_{00} & v^T \\ v & Q^{-1} \end{pmatrix} = \begin{pmatrix} a_{00}^{-1} + a_{00}^{-2}a^T Q^{-1} a & -a_{00}^{-1}a^T Q^{-1} \\ -a_{00}^{-1}Q^{-1}a & Q^{-1} \end{pmatrix}, \tag{14}$$

where $Q = D + bb^T - a_{00}^{-1}aa^T$, so that Q^{-1} satisfies (25) in section 6.1.

Savin, Wimmer and Witkovsky (2003) considered the interval estimation of the common mean by using (6). However, the information matrix for restricted likelihood in the Appendix to their paper seems to be in error.

To use (7) one needs the form of matrices Λ and H. Let p-dimensional vectors β, e have coordinates $\omega_i^2 \Phi_1^{-1/2}$ and 1 respectively. We also define the diagonal matrix Ψ to have nonzero elements $\psi_{ii} = \omega_i^3 \Phi^{-1}, i = 1, \ldots, p$, and

denote by I the $p \times p$ identity matrix. Then a direct calculation detailed in section 6.2 shows that

$$\Lambda = -\frac{1}{2}H = \begin{pmatrix} e^T \\ I \end{pmatrix} (\Psi - \beta\beta^T) \begin{pmatrix} e & I \end{pmatrix}, \qquad (15)$$

so that

$$\operatorname{tr}(\mathcal{V}\Lambda) = \operatorname{tr}\left((\Psi - \beta\beta^T)(v_{00}ee^T + ev^T + ve^T + Q^{-1})\right)$$

$$\approx \frac{1}{\Phi}\left[(S_3 - S_2^2)v_{00} + 2\sum \omega_i^2(\omega_i - S_2)v_i + \sum_i \frac{\omega_i^3(1 - \omega_i)}{d_{ii}}\right], \qquad (16)$$

as $\operatorname{tr}((\Psi - \beta\beta^T)Q^{-1})$ is numerically very close to $\operatorname{tr}((\Psi - \beta\beta^T)D^{-1})$. While $\mathcal{V}$ has full rank $p + 1$, the matrices Λ and $\Psi - \beta\beta^T$ have rank $p - 1$.

The true REML estimator of σ^2 is $\tilde{\sigma}^2 = \max(0, \tilde{\sigma}_{RL}^2)$, where $\tilde{\sigma}_{RL}^2$ is the (unbiased) solution of the restricted likelihood equation, $\partial RL/\partial\sigma^2 = 0$. Its approximate variance $v_{00} = a_{00}^{-2}(a_{00} + aQ^{-1}a^T)$ can be evaluated from the formulas of section 6.1.

For example, when $p = 2$,

$$\tilde{\sigma}_{RL}^2 = \frac{(x_1 - x_2)^2 - s_1^2 - s_2^2}{2}$$

has the variance $[(2\sigma^2 + \sigma_1^2 + \sigma_2^2)^2 + \sigma_1^4/\nu_1 + \sigma_2^4/\nu_2]/2$ which happens to coincide with v_{00}.

The largest discrepancy between $\tilde{\sigma}_{RL}^2$ and its truncated version $\tilde{\sigma}^2$ occurs when $\sigma^2 = 0$. Then $0.5 \leq P(\tilde{\sigma}_{RL}^2 \neq \tilde{\sigma}^2) = P(\chi^2(p - 1) \leq p - 1)$. This probability takes its largest value, $2\Phi(1) - 1 = 0.6827..$, when $p = 2$. The formulas for $E\tilde{\sigma}^2$ and $E(\tilde{\sigma}^2 - \sigma^2)^2$ can be found in section 6.3.

Figure 1 shows the plots of approximations to the variance of the REML estimator which was obtained from 10000 simulations for each of the values of $\sigma^2 = 0 : 0.2 : 5$ when $p = 2, n_1 = 5, n_2 = 3, \sigma_1^2 = 0.1, \sigma_2^2 = 0.5$. This figure and results of other simulations show that the approximation (6) is not adequate in our problem as it substantially overestimates $\operatorname{Var}(\tilde{x})$ for small σ^2.

3 Coverage factors

According to (16) if a truncated estimator $\tilde{\sigma}^2$ is employed,

$$\widetilde{Var}(\tilde{x}) = \tilde{\Phi} - \tilde{\Upsilon}\tilde{S}_2 + 2\operatorname{tr}(\tilde{\mathcal{V}}\tilde{\Lambda}) \approx \tilde{\Phi} - \tilde{\Upsilon}\tilde{S}_2 \qquad (17)$$
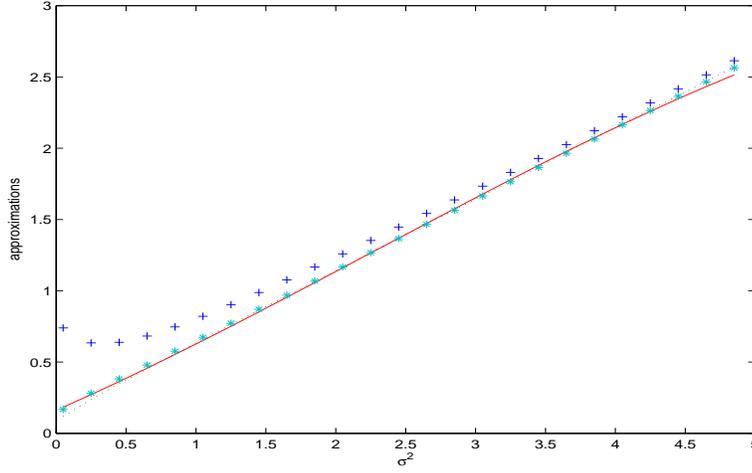
Figure 1: Plot of the variance of REML estimator (continuous line), and its approximations (6) (line marked by $+$), (7) (line marked by $*$), and of $\tilde{\Phi}$ (dotted line) when $p = 2$.

$$+\frac{2}{\tilde{\Phi}}\Big[(\tilde{S}_3-\tilde{S}_2^2)E(\tilde{\sigma}^2-\sigma^2)^2+2\sum\tilde{\omega}_i^2(\tilde{\omega}_i-\tilde{S}_2)E\tilde{\sigma}^2(\tilde{\sigma}_i^2-\sigma_i^2)+\sum_i\frac{2\tilde{\omega}_i^3(1-\tilde{\omega}_i)\tilde{\sigma}_i^4}{\nu_i}\Big],$$

with $E\tilde{\sigma}^2, E(\tilde{\sigma}^2-\sigma^2)(\tilde{\sigma}_i^2-\sigma_i^2)$ and $E(\tilde{\sigma}^2-\sigma^2)^2$ derived from formulas (28), (29), (30) in section 6.3, and $\tilde{S}_k=\sum_i\tilde{\omega}_i^k, \quad k=1,2,\ldots$.

For large σ^2 the approximation (26) there shows that

$$\widetilde{Var}(\tilde{x}) \sim \frac{\sum(\sigma^2+\sigma_i^2)\chi^2(p-1)}{p(p-1)},$$

and the pivot $(\tilde{x}-\mu)/\sqrt{\widetilde{Var}(\tilde{x})}$ is approximable by a $t(p-1)$-distribution. This approximation leads to the $(1-\alpha)-$approximate confidence intervals,

$$\tilde{x} \pm t_{1-\alpha/2}(p-1)\sqrt{\widetilde{Var}(\tilde{x})}, \tag{18}$$

where $\widetilde{Var}(\tilde{x})$ is found from (17). This interval will be compared to

$$\tilde{x} \pm t_{1-\alpha/2}(p-1)\sqrt{\tilde{\Phi}}. \tag{19}$$

In addition to (19) there are confidence intervals for μ which do not use Φ directly. One of them is based on Horn, Horn and Duncan (1975) procedure,

$$\tilde{x} \pm t_{1-\alpha/2}(p-1)\sqrt{\sum_i\frac{\tilde{\omega}_i^2(x_i-\tilde{x})^2}{1-\tilde{\omega}_i}}. \tag{20}$$

7

Another interval discussed in Sec 7.3.4 of Hartung, Knapp and Sinha (2008) is

$$\tilde{x} \pm t_{1-\alpha/2}(p-1)\sqrt{\sum_i \frac{\tilde{\omega}_i(x_i - \tilde{x})^2}{p-1}}. \tag{21}$$

It is based on the so-called external consistency estimator of the variance (Dietrich, 1991). When $p = 2$, (20) and (21) coincide both having the half-width $t_{1-\alpha/2}(1)\ \sqrt{\tilde{\omega}_1\tilde{\omega}_2}\ |x_1 - x_2|$.

Kenward and Roger (1997) proposed to use the Satterthwaite approximation for the pivotal quantity $(\tilde{x}-\mu)/\sqrt{\widetilde{Var}(\tilde{x})}$. In our case it means the Student distribution $t(m)$ whose degrees of freedom m are to be estimated from the data. However, the estimator suggested by these authors always exceeds 4, and does not provide a good approximation for small p. Also the assumption of independence between $\tilde{x}$ and $\widetilde{Var}(\tilde{x})$ used by these authors according to (4) leads to the formula, $E(\tilde{x} - \mu)^2/\widetilde{Var}(\tilde{x}) = E(\tilde{x} - \mu)^2 E[\widetilde{Var}(\tilde{x})]^{-1} \approx (\mathrm{tr}(\mathcal{W}\Lambda) + \Phi)/(2\mathrm{tr}(\mathcal{W}\Lambda) + \Phi) < 1$, which cannot be equal to the variance of a $t(m)$.

4   Numerical Results

We report here some results of the numerical comparison of the REML estimator based confidence interval (18) with the intervals (19), (20) and (21). In our Monte Carlo simulation study for $p = 2, 3, 5, 9$, we used randomly chosen sample sizes n_i with the uniform distribution over integers from 2 to 12 The error variances σ_i^2 were taken to have a scaled χ^2-distribution, so that $E\sigma_i^2 = 1$. Figures 2-5 display the coverage probability of these intervals with a nominal confidence coefficient of 95% which is reported as a function of $\sigma^2 = 0 : 0.2 : 5$. Both intervals (20) and (21) have lower than stated confidence level. The confidence intervals based on the DerSimonian-Laird estimator (not shown here) sustain the nominal confidence coefficient much better. The interval (21) outperformed (20) in our simulations. For $p \geq 3$, the interval (19) is not adequate when σ^2, is large, when it is small the interval (18) is too wide and too conservative as is seen from these Figures. For larger values of $p(p = 10, 15, 20, 30, 50)$, the behavior of (21) was studied in Sidik and Jonkman (2002) and the two intervals (20) and (21) are compared by Sidik and Jonkman (2006) who describe (20) as robust and recommend it.

The REML estimator was computed via its R-language implementation (through the *lme* function from the *nlme* library). The *intervals* function

with fixed effects also provides approximate confidence interval for μ which is too short for small/moderate p, but which can be adjusted by changing the normal quantile to that of a $t(p-1)$-distribution as in (19). The parameter estimates provided in the summary of *lme* are ratios $\tau_2^2/\tau_1^2, \ldots, \tau_p^2/\tau_1^2$, with the values τ_1^2 and σ^2 provided by *VarCorr* function. For multimodal or flat restricted likelihood functions convergence of the lme algorithm is problematic. Additional difficulty is caused by its non-convergence in cases where $\tilde{\sigma}^2 \approx 0$, in particular when for some i, s_i^2 is very small. In the latter case $\tilde{\sigma}_i^2 \approx 0$, and $\tilde{\sigma}^2 = 0$, so that $\tilde{\Phi} \approx 0$, and all intervals (19), (20) and (21) shrink towards $\tilde{x}$. For these reasons the results were cross-checked by the iterative algorithm in Rukhin (2011).

Figure 5 depicts a quite nonlinnear $q - q$ plot of pivotal quantity $(\tilde{x} - \mu)/\sqrt{\widetilde{Var}(\tilde{x})}$ against $t(p-1)$ when $p = 7$ and $\sigma^2 = 0.2$ obtained from $50,000$ runs. Simulations also show that $\mathrm{tr}\left(\tilde{\mathcal{V}}\tilde{\Lambda}\right)$ overestimates $E(\tilde{x} - \tilde{\mu})^2$ so that numerical accuracy of (4) is questionable.
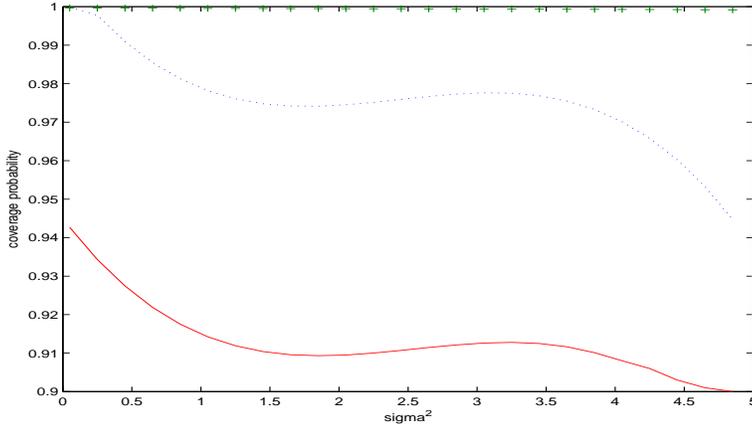
Figure 2: Plot of coverage probabilities of the confidence intervals (18) (dotted line), (19) (line marked by +), and (20), (21) (continuous line), when $p = 2$.

5 Conclusions

All simulation results suggest that in the heterogeneous setting of random effects meta-analysis the approximation (7) for $p \leq 10$ does not lead to good confidence intervals. Indeed one of its features is that since $\tilde{\Upsilon}$ is positive,

9
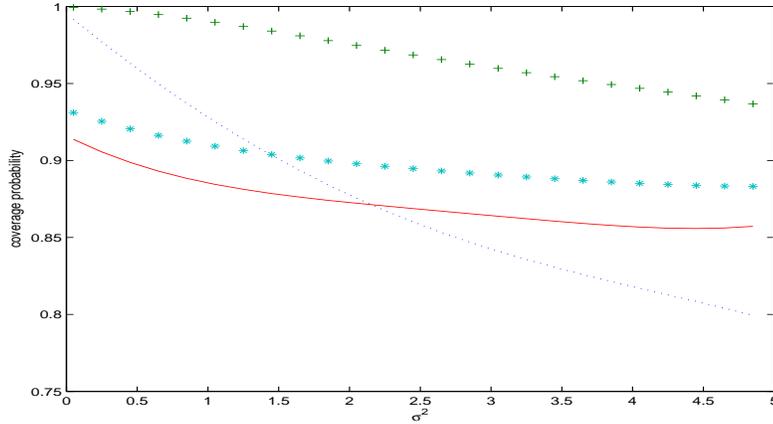
Figure 3: Plot of coverage probabilities of four confidence intervals with (21) marked by $*$, and (20) (continuous line), when $p = 3$ (other intervals have designations of lines are the same as in Figure 2).

two correction terms in (7) are of different signs, and both of them can be substantial with possibly negative sum. When σ^2 is small, these terms tend to cancel one another. When σ^2 is large, they are small and can be neglected. However, omitting them altogether as in (19) results in poor coverage probability for large σ^2 Intuitively, the biased estimator $\tilde{\sigma}^2$ has a patently non-normal distribution with a large mean squared error. See Figure 6 which shows the histogram of the distribution of $\tilde{\sigma}^2/\sigma^2$ when $p = 5, \sigma^2 = 2.5$ under scenario of section 4. Therefore for small/medium p, the higher order terms in Taylor's formula (7) cannot be discarded, and its accuracy is poor. The interval (21) based on the external consistency estimator of the variance performed the best in our simulations although (20) was always close.

6 Appendix

6.1 Covariance matrix for restricted maximum likelihood procedure

Since

$$E(x_i - \tilde{\mu})(x_j - \tilde{\mu}) = \delta_{ij}(\sigma^2 + \sigma_i^2) - \Phi, \qquad (22)$$

$$\tilde{\mu}_0' = \frac{\partial}{\partial \sigma^2}\tilde{\mu} = -\sum_i \frac{\omega_i^2(x_i - \tilde{\mu})}{\Phi}, \qquad (23)$$

10

and
$$\tilde{\mu}_i' = \frac{\partial}{\partial \sigma_i^2} \tilde{\mu} = -\frac{\omega_i^2 (x_i - \tilde{\mu})}{\Phi}, \tag{24}$$

one has
$$E(x_i - \tilde{\mu})\tilde{\mu}_0' = -\omega_i + S_2,$$

and
$$E(x_j - \tilde{\mu})\tilde{\mu}_i' = -\delta_{ij}\omega_i + \omega_i^2.$$

By differentiating RL one gets,

$$\frac{\partial}{\partial \sigma^2} RL = \frac{1}{2\Phi^2} \left[\sum_i \omega_i^2 (x_i - \tilde{\mu})^2 - 1 + S_2 \right],$$

$$-\frac{\partial^2}{\partial \sigma^4} RL = \frac{1}{\Phi^2} \left[\sum_i \omega_i^2 (x_i - \tilde{\mu})\tilde{\mu}_0' + \sum_i \frac{\omega_i^3 (x_i - \tilde{\mu})^2}{\Phi} - \frac{1}{2}S_2 + S_3 - \frac{S_2^2}{2} \right],$$

and

$$-\frac{\partial^2}{\partial \sigma^2 \partial \sigma_i^2} RL = \frac{1}{\Phi^2} \left[\sum_j \omega_j^2 (x_j - \tilde{\mu})\tilde{\mu}_i' + \frac{\omega_i^2 (x_i - \tilde{\mu})^2}{\Phi} - \frac{\omega_i^2}{2} + \frac{\omega_i^3}{\Phi} - \frac{\omega_i^2 S_2}{2} \right],$$

so that
$$\mathcal{J}_{00} = -E \frac{\partial^2}{\partial \sigma^4} RL = \frac{S_2 - 2S_3 + S_2^2}{2},$$

$$\mathcal{J}_{0i} = -E \frac{\partial^2}{\partial \sigma^2 \partial \sigma_i^2} RL = \frac{\omega_i^2}{2\Phi^2} \left(1 + S_2 - 2\omega_i \right).$$

Similarly,

$$-\frac{\partial}{\partial \sigma_i^2} RL = \frac{1}{2} \left[-\frac{\omega_i^2 (x_i - \tilde{\mu})^2}{\Phi^2} + \frac{\omega_i (1 - \omega_i)}{\Phi} + \nu_i \left(\frac{1}{\sigma_i^2} - \frac{s_i^2}{\sigma_i^4} \right) \right],$$

$$-\frac{\partial^2}{\partial \sigma_i^4} RL = \frac{\omega_i^2 (x_i - \tilde{\mu})\tilde{\mu}_i'}{\Phi^2} + \frac{\omega_i^2 (x_i - \tilde{\mu})^2}{\Phi^3} - \frac{\omega_i^2 (1 - \omega_i)^2}{2\Phi^2} - \nu_i \left(\frac{1}{2\sigma_i^4} - \frac{s_i^2}{\sigma_i^6} \right),$$

and for $i \neq j$

$$-\frac{\partial^2}{\partial \sigma_i^2 \partial \sigma_j^2} RL = -\frac{\omega_i \omega_j (x_i - \tilde{\mu})(x_j - \tilde{\mu})}{\Phi^3} - \frac{\omega_i^2 \omega_j^2}{2\Phi^2}.$$

It follows that

$$\mathcal{J}_{ii} = -E \frac{\partial^2}{\partial \sigma_i^4} RL = \frac{\nu_i}{2\sigma_i^4} + \frac{\omega_i^2 (1 - \omega_i)^2}{2\Phi^2},$$

11

$$\mathcal{J}_{ij} = -E\frac{\partial^2}{\partial\sigma_i^2\partial\sigma_j^2}RL = \frac{\omega_i^2\omega_j^2}{2\Phi^2},$$

if $1 \le i \ne j \le p$. The representation (9) of the information matrix $\mathcal{J}$ follows, and the form of its inverse (14) can be derived from standard formulas (e.g. Theorem 18.2.8 in Harville, 1997).

To find the matrix $Q^{-1} = (D + bb^T - a_{00}^{-1}aa^T)^{-1}$ in (14), we put $c = D^{-1/2}b, d = D^{-1/2}a/\sqrt{a_{00}}$, so that

$$Q = D^{1/2}(I + cc^T - dd^T)D^{1/2}.$$

Thus,

$$D^{1/2}Q^{-1}D^{1/2} = (I + cc^T - dd^T)^{-1} = \left[I + (c,d)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} c^T \\ d^T \end{pmatrix}\right]^{-1}$$

$$= I - \frac{1}{\Delta}\left[(d^Td - 1)cc^T - (c^Td)(dc^T + cd^T) + (c^Tc + 1)dd^T\right],$$

where $\Delta = (c^Tc + 1)(d^Td - 1) - (c^Td)^2$ is the determinant of the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} c^T \\ d^T \end{pmatrix}(c,d)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} c^Tc + 1 & -c^Td \\ -c^Td & d^Td - 1 \end{pmatrix}.$$

Therefore,

$$Q^{-1} = D^{-1} - \frac{1}{a_{00}\Delta}\Big[(a^TD^{-1}a - a_{00})D^{-1}bb^TD^{-1} \tag{25}$$

$$-(b^TD^{-1}a)D^{-1}(ab^T + ba^T)D^{-1} + (b^TD^{-1}b + 1)D^{-1}aa^TD^{-1}\Big],$$

$$a^TQ^{-1}a = a^TD^{-1}a + \frac{(a^TD^{-1}a)^2(b^TD^{-1}b + 1) - (b^TD^{-1}a)^2(a^TD^{-1}a + a_{00})}{a_{00}|\Delta|}$$

$$= \frac{a^TD^{-1}a + (a^TD^{-1}a)(b^TD^{-1}b) - (b^TD^{-1}a)^2}{|\Delta|},$$

as $\Delta = (b^TD^{-1}b + 1)(a^TD^{-1}a/a_{00} - 1) - (b^TD^{-1}a)^2/a_{00} < 0$. Indeed since $d_{ii} > a_i$, $a^TD^{-1}a = \sum a_i^2 d_{ii}^{-1} \le \sum a_i = (S_2 - 2S_3 + S_2^2)\Phi^{-2} = a_{00}$.

For large σ^2,

$$v_{00} = \frac{a_{00} + aQ^{-1}a^T}{a_{00}^2}$$

$$= \frac{2\sigma^4 p(p - 1) + 4\sigma^2(p - 1)\sum\sigma_i^2 - 6(p - 2)\sum\sigma_i^4 + 2(4 - 7p^{-1})(\sum\sigma_i^2)^2}{p(p - 1)^2}$$

12

$$+ \sum_i \frac{2\sigma_i^4}{p^2 \nu_i} + O\left(\frac{1}{\sigma^2}\right),$$

and

$$v_i = -\frac{2\sigma_i^4}{p \nu_i} + O\left(\frac{1}{\sigma^2}\right).$$

These formulas suggest the following approximations in distribution,

$$\tilde{\sigma}_{RL}^2 \sim \frac{\sum(\sigma^2 + \sigma_i^2)\chi_{p-1}^2}{p(p-1)} - \frac{\sum \tilde{\sigma}_i^2}{p}, \quad \tilde{\sigma}_i^2 \sim \sigma_i^2 \frac{\chi^2(\nu_i)}{\nu_i}, \qquad (26)$$

which are used in section 6.3 to derive formulas for $E\tilde{\sigma}^2$, $E\tilde{\sigma}^2(\tilde{\sigma}_i^2 - \sigma_i^2)$, and $E(\tilde{\sigma}^2 - \sigma^2)^2$ assuming approximate independence of $\tilde{\sigma}_i^2$.

The correction term in (16) has the asymptotic expansion of the form

$$\mathrm{tr}(\mathcal{V}\Lambda) = \frac{(p-1)}{p^3 \sigma^2}\left[p\sum \sigma_i^4 - \left(\sum \sigma_i^2\right)^2 + \sum \frac{2\sigma_i^4}{\nu_i}\right] + O\left(\frac{1}{\sigma^4}\right),$$

so that

$$Var(\tilde{x}) = \frac{\sigma^2}{p}\left[1 + \frac{1}{p\sigma^2}\sum \sigma_i^2\right.$$

$$\left. + \frac{(2p-3)}{p^2\sigma^4}\left(p\sum \sigma_i^4 - (\sum \sigma_i^2)^2\right) + \frac{4(p-1)}{p^2\sigma^4}\sum \frac{\sigma_i^4}{\nu_i}\right] + O\left(\frac{1}{\sigma^4}\right).$$

6.2 Covariance matrix of $\tilde{\mu}'$

Formulas (22), (23) and (24) can be used to find the elements of matrix Λ. Indeed

$$E(\tilde{\mu}_0')^2 = \frac{1}{\Phi^2}\sum_{i,j} E\omega_i\omega_j(x_i - \tilde{\mu})(x_j - \tilde{\mu}).$$

Similarly,

$$E\tilde{\mu}_0'\tilde{\mu}_i' = -\frac{\omega_i^2 E\tilde{\mu}_0'(x_j - \tilde{\mu})}{\Phi}.$$

Derivation of other elements of matrix Λ is straightforward, and the form (15) follows. The rank of this matrix is $p - 1$, as $\tilde{\mu}_0' = \sum \tilde{\mu}_i'$, and $\sum(\sigma^2 + \sigma_i^2)\tilde{\mu}_i' = 0$.

By differentiating $\Phi = \Phi(\sigma^2, \sigma_1^2, \ldots, \sigma_p^2)$ twice, we see that its Hessian H coincides with $-\Lambda/2$. This means that the function Φ is concave, but it is not strictly concave as rank $(H) = \mathrm{rank}(\Lambda) = p - 1 < p + 1$.

13

6.3 Estimated covariance matrix

According to the approximation (26), the expected value of $\tilde{\sigma}^2$ has the form with $z^2 = (p-1)(\sum \tilde{\sigma}_i^2)/[2\sum(\sigma^2 + \sigma_i^2)]$,

$$E\tilde{\sigma}^2 = \frac{2\sum(\sigma^2 + \sigma_i^2)}{p(p-1)\Gamma((p-1)/2)} E \int_{z^2}^{\infty} (u - z^2)u^{(p-1)/2-1}e^{-u} \, du.$$

For example, if $p - 1 = 2k, k \geq 1$, is an even integer,

$$E\tilde{\sigma}^2 = \frac{2\sum(\sigma^2 + \sigma_i^2)}{p(p-1)} \sum_{j=0}^{k-1} \frac{(k-j)}{j!} Ez^{2j}e^{-z^2}. \tag{27}$$

For any non-negative integer j

$$Ez^{2j}e^{-z^2} = (-1)^j \frac{d^j}{dt^j} Ee^{-tz^2} \big|_{t=1},$$

and for any positive t,

$$Ee^{-tz^2} = \prod_i \left(1 + \frac{t(p-1)\sigma_i^2}{\nu_i \sum(\sigma^2 + \sigma_i^2)}\right)^{-\nu_i/2},$$

so that the expected values in (27) can be readily evaluated. Similarly, in this case

$$E\tilde{\sigma}^4 = \frac{[2\sum(\sigma^2 + \sigma_i^2)]^2}{[p(p-1)]^2\Gamma((p-1)/2)} E \int_{z^2}^{\infty} (u - z^2)^2 e^{-u} u^{(p-1)/2-1} \, du$$

$$= \frac{[2\sum(\sigma^2 + \sigma_i^2)]^2}{[p(p-1)]^2} \sum_{j=0}^{k-1} \frac{(k+1-j)(k-j)}{j!} Ez^{2j}e^{-z^2}.$$

Somewhat more complicated formulas exist for any p. When $(p-1)/2 = \ell + 1/2$ with an integer $\ell \geq 0$, the identity,

$$\frac{e^{-u}}{\sqrt{u}} = \frac{1}{\Gamma(1/2)} \int_1^{\infty} \frac{e^{-tu} \, dt}{\sqrt{t-1}},$$

can be used to express the mean squared error of the estimator $\tilde{\sigma}^2$ as a linear combination of integrals involving terms $Ez^{2j}e^{-tz^2}, j = 0, 1, \ldots, \ell$. For example,

$$E\tilde{\sigma}^2 = \frac{2\sum(\sigma^2 + \sigma_i^2)}{p(p-1)\Gamma((p-1)/2)} E \int_{z^2}^{\infty} (u - z^2)u^\ell u^{-1/2}e^{-u} \, du$$

$$= \frac{2\sum(\sigma^2 + \sigma_i^2)}{p(p-1)\Gamma((p-1)/2)\Gamma(1/2)} E \int_1^\infty \int_{z^2}^\infty \frac{(u - z^2)u^\ell e^{-tu} \, du \, dt}{\sqrt{t-1}}$$

$$= \frac{\sum(\sigma^2 + \sigma_i^2)\Gamma(\ell+1)}{p\Gamma((p+1)/2)\Gamma(1/2)} \sum_{j=0}^\ell \frac{(\ell+1-j)}{j!} \int_1^\infty \frac{t^{j-\ell-2} \, dt}{\sqrt{t-1}} Ez^{2j}e^{-tz^2}, \qquad (28)$$

$$E\tilde{\sigma}^2(\tilde{\sigma}_k^2 - \sigma_k^2)$$

$$= \frac{\sum(\sigma^2 + \sigma_i^2)\Gamma(\ell+1)}{p\Gamma((p+1)/2)\Gamma(1/2)} \sum_{j=0}^\ell \frac{(\ell+1-j)}{j!} \int_1^\infty \frac{t^{j-\ell-2} \, dt}{\sqrt{t-1}} Ez^{2j}(\tilde{\sigma}_k^2 - \sigma_k^2)e^{-tz^2},$$

$$(29)$$

and

$$E\tilde{\sigma}^4 = \frac{2[\sum(\sigma^2 + \sigma_i^2)]^2\Gamma(\ell+1)}{p^2(p-1)\Gamma((p+1)/2)\Gamma(1/2)} \qquad (30)$$

$$\times \sum_{j=0}^\ell \frac{(\ell+2-j)(\ell+1-j)}{j!} \int_1^\infty \frac{t^{j-\ell-3} \, dt}{\sqrt{t-1}} Ez^{2j}e^{-tz^2}.$$

These integrals are easy to evaluate by standard numerical integration formulas, e.g., via the trapezoid method. There is a quadrature formula (Harper, 1962) which can be used after a transformation of variables $t = 1 + y^2$. Similar integrals appear as the posterior distribution in the Bayesian setting of the Behrens-Fisher problem.

For large σ^2,

$$\Upsilon \approx \frac{E(\sum \tilde{\sigma}_i^2)^{(p+1)/2}}{\Gamma((p+3)/2)[2\sum(\sigma^2 + \sigma_i^2)/(p-1)]^{(p-1)/2}},$$

and

$$v_{00} - E(\tilde{\sigma}^2 - \sigma^2)^2 \approx 2\sigma^2\Upsilon = O\left(\frac{1}{[\sum(\sigma^2 + \sigma_i^2)]^{(p-3)/2}}\right).$$

References

[1] Borenstein, M., Hedges, L., Higgins, J., and Rothstein H. (2009), *Introduction to Meta-Analysis*, Wiley, New York.

[2] Dietrich C. F., 1991. *Uncertainty, Calibration and Probability: The Statistics of Scientific and Industrial Measurement*, 2nd ed A. Hilger, London.

[3] Harper, W., 1962. Quadrature formulas for infinite integrals, *Mathematics of Computation*, 16, 170-175.

[4] Hartung, J., Knapp, G., and Sinha, B. K., 2008. *Statistical Meta-Analysis with Applications*, Wiley, New York.

[5] Harville, D. A., 1997. *Matrix Algebra from a Statistician's Perspective*, Springer, New York.

[6] Harville, D. A., 1985. Maximum likelihood approaches to variance components estimation and related problems, *Journal of the American Statistical Association*, 72, 320–340.

[7] Harville, D. A., and Jeske D. R., 1992. Mean squared error of estimation or prediction under a general linear model, *Journal of the American Statistical Association*, 87, 724–731.

[8] Horn, S. A., Horn, R. A., and Duncan, D. B., 1975. Estimating heteroscedastic variance in linear models, *Journal of the American Statistical Association*, 70, 380–385.

[9] Kackar, R., and Harville, D. A., 1984. Approximations for standard errors of estimators of fixed and random effects in mixed linear models, *Journal of the American Statistical Association*, 79, 853–862.

[10] Kenward, M. G., and Roger, J. H., 1997. Small sample inference for fixed effects from restricted maximum likelihood, *Biometrics*, 53, 983–997.

[11] Kenward, M. G., and Roger, J. H., 2009. An improved approximation to the precision of fixed effects from restricted maximum likelihood, *Computational Statistics&Data Analysis*, 53, 2583-2595.

[12] Rukhin, A. L., 2011. Maximum likelihood and restricted maximum likelihood solutions in multiple-method studies, *Journal of Research of the National Institute of Standards and Technology*, 116, 539-556.

[13] Savin, A., Wimmer, G., and Witkovsky, V., 2003. On Kenward-Roger confidence intervals for common mean in interlaboratory trials, *Measurement Science Review*, 3, 53-56.

[14] Sidik, K., and Jonkman, J., 2002. A simple confidence interval for meta-analysis, *Statistics in Medicine*, 21, 3153-3159.

[15] Sidik, K., and Jonkman, J., 2006. Robust variance estimation for random effects meta-analysis, *Computational Statistics&Data Analysis*, 50, 3681-3701.
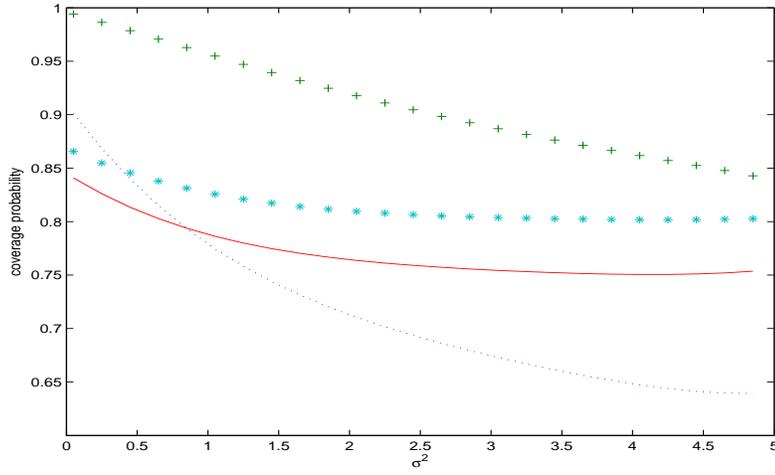
Figure 4: Plot of coverage probabilities of four confidence intervals when $p = 5$ (designations of lines are the same as in Figure 3).
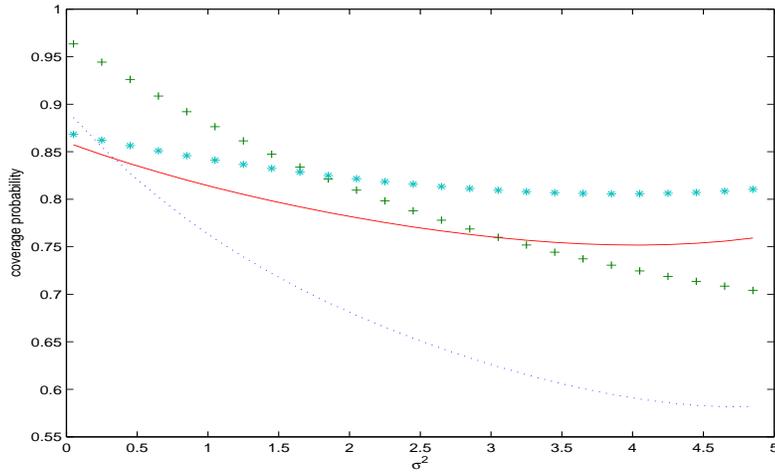
Figure 5: Plot of coverage probabilities of four confidence intervals when $p = 9$.

17
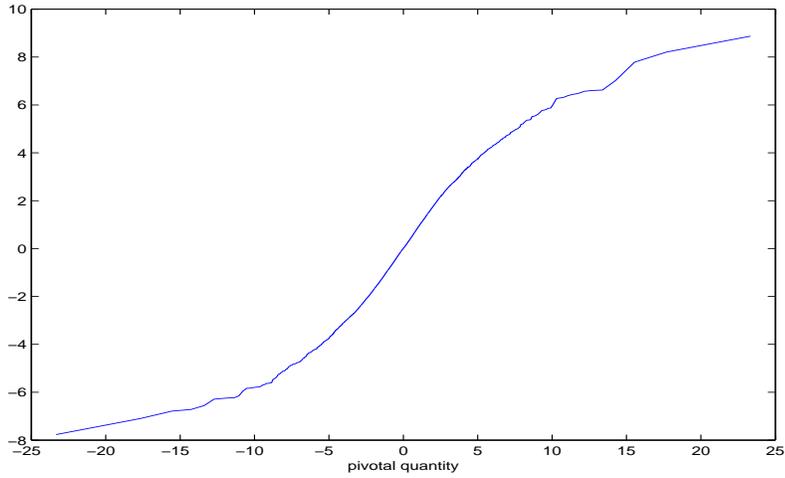
Figure 6: The q-q plot of $(\tilde{x} - \mu)/\sqrt{\widetilde{Var(\tilde{x})}}$ when $p = 5, \sigma^2 = 2.5$.
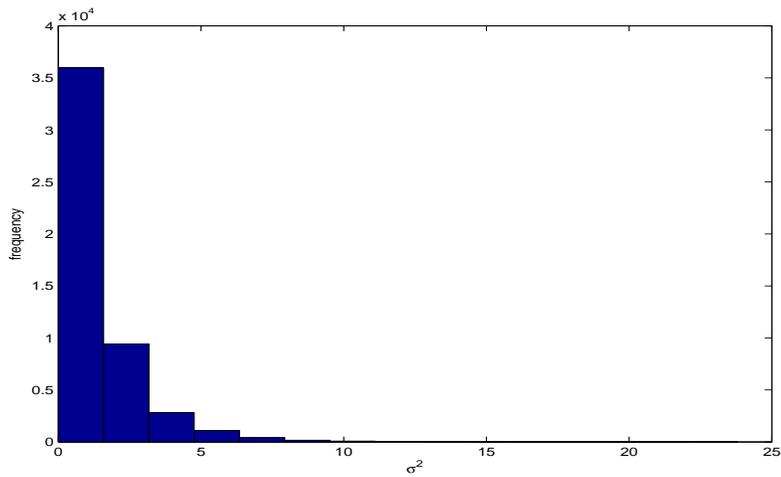
Figure 7: Histogram of $\tilde{\sigma}^2/\sigma^2$ when $p = 5, \sigma^2 = 2.5$.

18